Appendix A

Summary of Peer Review Team Comments and Study Authors' Responses

The Peer Review Team consisted of the following individuals.

- Mr. Leon Billings
- Mr. Tom Brosnan, National Oceanic and Atmospheric Administration
- Mr. Michael Cook, U.S. Environmental Protection Agency
- Mr. John Dunn, U.S. Environmental Protection Agency
- Dr. Mohammad Habibian, Washington Suburban Sanitation Commission
- Dr. Leo Hetling, Public Health and Environmental Engineering, New York State Department of Environmental Conservation (retired)
- Dr. Russell Isaacs, Massachusetts Department of Environmental Protection
- Dr. Norbert Jaworski, U.S. Environmental Protection Agency (retired)
- Dr. William Jobin, Blue Nile Associates
- Mr. Ken Kirk, American Metropolitan Sewerage Association
- Mr. John Kosco, U.S. Environmental Protection Agency
- Mr. Rich Kuhlman, U.S. Environmental Protection Agency
- Mr. Joseph Lagnese
- Ms. Jessica Landman, National Resource Defense Council
- Mr. Kris Lindstrom, K.P. Lindstrom, Inc.
- Mr. Ronald Linsky, National Water Research Institute
- Dr. Berry Lyons, University of Alabama
- Dr. Alan Mearns, National Oceanic and Atmospheric Administration
- Dr. Daniel Okun, University of North Carolina
- Mr. Steve Parker, National Research Council
- Mr. Richard Smith, U.S. Geological Survey
- Mr. Phill Taylor, U.S. Environmental Protection Agency and Tetra Tech, Inc. (retired)

- Dr. Red Wolman, Johns Hopkins University

EPA Instructions to Peer Review Team for Evaluation of Draft Report Dated October 18, 1998

- Feedback on accuracy and historical context of statements in the report.
- Evaluation of the reliability of the statistical techniques used to document "before and after" trends in dissolved oxygen.
- Have we overlooked any significant work in the literature relevant to the study?
- Have we missed anything in interpretations of "before and after" data in relation to the historical context of water pollution control activities in the United States?
- Has the study met the stated objective of identifying national-scale progress in water quality achieved as a result of EPA's investment in water pollution control infrastructure?

The following key issues were identified in the responses received from 21 members of the Peer Review Team. For each key issue, a summary of the major points raised by the members of the Peer Review Team is presented along with our responses to the issue and where the relevant information is presented in the final report.

Key Issues Identified by Peer Review Team

1. Target audience
2. Title of final report and objectives of the study
3. Executive Summary
4. Use of oxygen as key indicator for "before and after" trends
5. Availability of monitoring data in STORET
6. How representative is oxygen data used for "before and after" trends?
7. Statistical methods
8. Geographic representation of case study sites
9. Cost versus benefits analysis
10. Editing for final report

Issue 1: Target Audience

Reviewers:

Linsky, Okun, Wolman, Hetling, Jobin, Isaacs, Parker, Billings , and Habibian

Comment summary:

Who is the audience for the report? How can the report provide useful information to policy makers, legislators, regulatory agencies, and the general public? What is the overall message? How can the report be structured to provide guidance for future water quality management programs? The report can be more than a

history; it can be an instrument for beneficial change. The report needs to clearly articulate the interrelatedness of the CWA and SDWA for water pollution control and water quality management in relation to both public health (drinking water) and ecological quality of rivers and streams. The CWA and the secondary treatment requirement have demonstrated some water quality success over 25 years. Future and continued successes for the next 25 years, however, are not a given unless national policies are based on a sustainable strategy driven by citizen stakeholder groups (Jobin). O&M costs are high and it is very important (but not politically highly visible) to maintain levels of funding as are replacement costs of obsolete POTWs. Jobin discusses three "traps" that have hampered long-term sustainability of past water pollution control efforts. He points out the need to present strong conclusions as well as recommendations to guide future efforts.

Response:

A discussion on the report audience has been included in Chapter 1 of the document. The primary audience is the technical scientists and engineers who try to evaluate cause-effect relationships of pollutant load and ambient water quality responses. The secondary audience is Congress, regulatory/policy professionals, and the informed public, who need to understand that a major public works program (the CWA Construction Grants and CWSRF programs) accomplished what it was designed to do—reduce BOD effluent loads from municipal and industrial sources and improve dissolved oxygen in many previously degraded waterways of the Nation. These same groups also need to understand that water pollution control efforts never end. The 1972 CWA did not "solve" the problem; in fact, waste materials are generated continuously and effluent removal efficiencies must increase in the future to compensate for population growth. Planning for O&M expenditures, as well as capital expenditures for replacement of obsolete facilities and upgrades to maintain adequate levels/efficiency of wastewater removal is an ongoing requirement. Chapter 2 and the Executive Summary include a projection analysis that demonstrates that many of the gains in national water quality improvements may be lost if future wastewater infrastructure capacity does not keep pace with expected urban population growth.

Issue 2: Title of Final Report and Objectives of Study

Reviewers:

Landman, Hetling, Mearns, Kosco

Comment summary:

The title of the report, *Progress in Water Quality: An Evaluation of the Benefits of the 1972 Clean Water Act*, is too broad and implies too large a scope of study (i.e., we address all issues of the CWA). The title needs to be changed to be more representative of the data presented in the study. The authors need to state more clearly up-front that the study was designed to evaluate the effectiveness of investments in POTW upgrades on improving oxygen levels in previously degraded waterways and nothing more. The caveat for the analysis is that POTW sources are only one component of many possible sources of oxygen-demanding loads to waterways. It is important to get across the concept that improvements to POTWs alone are not sufficient to restore and maintain water quality as a national goal. The study has demonstrated that upgrades to POTWs had the expected result of improving oxygen levels in waterways once characterized by low levels. Can changes in oxygen be isolated to the impact of POTW inputs alone? In this study the important contribution is the findings of improvements in oxygen that are linked to POTW upgrades and investments—not the methodology.

Response:

The title has been changed to *Progress in Water Quality: An Evaluation of the National Investment in Municipal Wastewater Treatment*. The objectives of the study are now clearly stated in Chapter 1 and the

Executive Summary. Also, Chapter 2 includes a section that compares POTW sources of BOD loading to other major source categories (industrial point sources, CSOs, and urban and rural nonpoint sources) based on EPA's NWPCAM. Statements have been added to stress that continued improvement in the Nation's water quality conditions will require control of all major pollution sources, of which POTWs are only a portion. (POTWs contribute about 21 percent of all point and nonpoint BOD sources nationwide.) In this document both the findings and the data analysis methodology are given equal emphasis. A complete presentation of the methodology allows the scientist or statistician to assess the level of rigor of the analysis, as well as demonstrate the potential application of the methodology to other water pollution control performance measures.

Issue 3: Executive Summary

Reviewers:

Cook, Kirk, Landman, Dunn, Linsky, Brosnan, Kuhlman

Comment summary:

The Executive Summary is too long with too much defense of the methodology used for the analysis. Most people, especially policy, regulatory, and general public, will read only the Executive Summary and not care at all about (or understand) the technical details of how the analysis was performed. The key findings need to be made very clear in a concise summary that in turn can be boiled down to a press release of a few pages (which in turn can be boiled down to a "sound bite" and a "headline"). The Executive Summary needs a discussion of the status of water quality nationwide in relation to all sources (synopsis of most recent national water quality report to Congress?) The sequence of material presented in the Executive Summary should follow the sequence presented in the main report. One suggestion is to publish only the Executive Summary and put all the technical documentation of the main report on the EPA web site. The authors need to present strong conclusions (Isaacs, Hetling) as well as solid recommendations (Hetling).

Response:

The Executive Summary has been entirely rewritten as a high-quality, "stand-alone" document that presents the key findings of the study succinctly and in the same sequence as the main report. Also, the main body of the report has been reorganized and streamlined and includes an introductory chapter (Chapter 1, which is a road map to the rest of the report) as well as a summary and conclusions section at the end of each chapter.

Issue 4: Use of Dissolved Oxygen as Key Water Quality Indicator for "Before and After" Trends

Reviewers:

Hetling, Parker, Linsky, Jobin, Dunn, Smith, Jaworski

Comment summary:

Some reviewers, but not all, agreed that oxygen was most appropriate for questions posed for the study. One reviewer suggested that dissolved oxygen saturation should have been used as the indicator rather than dissolved oxygen concentration. The same reviewer also brought up a variety of statistical questions about the approach and the validity of the techniques used for the analysis of dissolved oxygen trends. The reviewers' consensus was that a more detailed rationale for the selection of dissolved oxygen as the key water quality indicator used for the study was needed in the final report. The reviewers also recommended that a better explanation of water quality standards for dissolved oxygen and the choice of 5 mg/L as a benchmark

concentration for comparison of "before and after" conditions be included in the final report. Some reviewers suggested combining the discussion of factors affecting dissolved oxygen in rivers with the discussion of pollutant loading.

Response:

Dissolved oxygen is a key chemical measure that has been used for many decades to characterize the overall health of aquatic ecosystems. High concentrations (~5 mg/L) greater than about 60 percent saturation levels are generally indicative of a healthy aquatic ecosystem whereas low concentrations (< 3 mg/L) less than about 40 percent saturation may be indicative of a stressed ecosystem. A large historical database with generally reliable measurements is available with records for a few waterways since the 1920s and 1930s. There is a well-understood causal relationship between municipal and industrial wastewater loading of oxidizable carbonaceous and nitrogenous materials (BOD), receiving waters' streamflow, and ambient concentrations of dissolved oxygen. Excessive depletion of dissolved oxygen as a result of poorly treated wastewater discharges was one of the major water pollution problems in many rivers and estuaries during the 1940s through the 1960s. The technology of secondary treatment, required as a minimum technology for municipal facilities by the 1972 CWA, is designed primarily to reduce the loading of BOD to improve dissolved oxygen conditions in streams, rivers, and estuaries.

This issue is discussed in the Executive Summary and in Chapters 2 and 3 in the final report.

Issue 5: Availability of Monitoring Data in STORET

Reviewers:

Wolman, Hetling, Kirk, Lagnese

Comment summary:

The fact that the "before and after" trends analysis was based on an apparently limited data set of only 246 catalog units out of a total of 1,666 catalog units with reaches impacted by point source discharges implies a significant problem with the availability and use of national-scale monitoring data from STORET for "performance evaluations" of point and nonpoint source pollution control measures. The apparent lack of a consistent and reliable water quality database on a national scale is a significant issue and needs to be discussed in more detail in the final report. The reviewers asked if the final report could present recommendations and conclusions about water quality monitoring programs that would provide guidance for future policy decisions or analysis efforts.

Response:

National-scale assessments of "performance evaluations" of the impact of water pollution control policies ideally should be based on a database large enough to provide a reliable sample of the effectiveness of implementation of regulatory policy. The objective of the study was a quantitative assessment of how much water quality has improved since the 1972 CWA. The purpose was to evaluate whether the national investment of $61 billion (current year dollars) that was targeted toward upgrading municipal wastewater treatment plants to secondary and better-than-secondary levels of treatment was, in fact, an effective regulatory policy for the Nation. As discussed under Issue 4, dissolved oxygen was selected as the key water quality indicator for the evaluation. To characterize long-term trends in oxygen, data sets needed to be compiled to represent "before" conditions prior to the 1972 CWA for comparison to "after" conditions. Persistent drought conditions were widespread in large areas of the northeast, middle Atlantic, and central United States during both the early 1960s (1962-1966) and the late 1980s (1987-1988). Early in the study, 1961-1965 and 1986-1990 were selected as the 5-year blocks of time to represent "before and after" conditions. The choice of 1961-1965 to represent conditions "before" the CWA was based, in part, on the availability of national-

scale dissolved oxygen data compiled from 1957-1965 by the FWQA (Gunnerson, 1965) from an early monitoring program funded under the 1956 Amendments to the Federal Water Pollution Control Act.

After the "before and after" analysis was completed, an inventory of the availability of oxygen and other water quality parameters in STORET from 1941-1995 was compiled. The limited availability of data for the "before" period constrained the sample size of the data sets that could be compiled for the "before and after" comparison. Nationwide, only 5,185 stations with 125,772 observations of oxygen were available for the 1961-1965 "before" period compared to 17,656 stations with 955,123 observations for the 1986-1990 "after" years. The largest abundance of stations were monitored after passage of the 1972 CWA with 34,052 stations and 749,125 observations recorded during 1971-1975. A considerably larger sample would most likely have been available for the "before and after" trend analysis if the 5-year period of 1966-1970 (16,972 stations) or 1971-1975 (34,052 stations) had been selected as the "before" period rather then the period of much more limited data availability, 1961-1965 (5,185 stations). Selection of 1961-1965 thus resulted in the analysis being based on only about one-third of the data available in 1966-1970 and only about one-sixth of the data available during the peak monitoring activity years of 1971-1975. Final selection of "before and after" data sets is based on comparable "dry" hydrologic characteristics as a "filter" for the analysis.

Even with the limited data set available for the "before and after" analysis, the results clearly document statistically significant improvements in dissolved oxygen with data sets aggregated over all spatial scales from the relatively small RF1 reach to the catalog unit and the major river basin. The methodology can be applied using 1966-1970 or 1971-1975 as the "before" period to enhance the robustness of the "before and after" analysis for dissolved oxygen trends. The methodology can also be applied for other water quality parameters for trend assessments of nutrients and TSS, for example, to evaluate the effectiveness of point and nonpoint source control programs on these parameters.

***Recommendations on monitoring programs**:*

- Sufficient federal funding to state and local government agencies should be made available for long time periods (5- to 10-year programs) to ensure support for continuous operation of a national-scale monitoring network and data collection efforts.
- Adequate federal funding must be made available to state and local governments to ensure the continuation of state-local-federal data submission and data management activities, including archives of historical water quality data, with STORET designated as the centrally accessible database for water quality data and information.
- State and local government agencies and university research groups receiving federal funds to support water quality monitoring programs should be required to submit all data subjected to QA/QC procedures in electronic format specified by EPA to STORET as the designated national database repository for water quality data. All water quality data records submitted to STORET should be required to be cross-referenced to geographic identifier codes defined for the Reach File and the National Hydrography Database.
- Interagency coordination for joint sharing of water quality data and submission of data to STORET should be established as standard operating procedure for all federal agencies involved in collecting water quality data for monitoring, research, or other programmatic purposes in freshwater and marine environments. Water quality data and information are collected by many federal agencies other than EPA; USGS, NOAA, NPS, USDA, USACE, U.S. Navy, and U.S. Coast Guard all collect water quality data for a variety of geohydrologic, limnological, and oceanographic purposes. Comprehensive national-scale assessments of the effectiveness of regulatory policies can be most successful only if all available data are shared (subject to national security clearance constraints imposed by the U.S. Navy) and pooled in a centrally accessible database designed to adhere to rigorous QA/QC procedures.

- Adequate federal funding must be made available to ensure the development of state-of-the-art software "tools" with up-to-date computer technology for the user community to be able to conduct credible statistical analyses to evaluate the status and trends of key water quality parameters. Raw data extracted from STORET and software "tools" should be made available on-line through EPA's Office of Water web site.

Issue 6: How Representative is Oxygen Data Used for "Before and After" Trends?

Reviewers:

Wolman, Hetling, Kirk, Lagnese, Parker, Lindstrom, Isaacs, Smith, Habibian

Comment summary:

Some reviewers thought that the "before and after" results were not particularly impressive as a basis to proclaim national-scale improvements in dissolved oxygen. Other reviewers did, however, think the "before and after" results were impressive and that the report presented solid evidence for documenting improvements in oxygen that were undoubtedly significant. Some reviewers pointed out that NPS loads, even in reaches impacted by point sources, probably confounded the analysis and contributed to "before and after" trends where big improvements were not identified. Some reviewers suggested that the final report present data showing (a) "before and after" oxygen trends for RF1 reaches impacted only by NPS under "dry" and "wet" conditions; (b) "before and after" oxygen data for RF1 reaches showing the worst degradation using the point source-impacted "dry" criteria; and (c) "before and after" BOD_5 trends to correlate with the oxygen data. The issues raised by the reviewers are twofold: (1) revise the report to present results to clearly document that RF1 reaches with "before and after" improvements probably represent a significant portion of the US population and municipal wastewater loads, and (2) show that the data filtering methodology for "dry" point source-dominated reaches accurately represents the "worst-case" conditions response of ambient oxygen to changes in waste loads.

Response:

Although the "before and after" evaluation of oxygen is based on data compiled from only 246 catalog units and 311 RF1 reaches, the waterways included in the analysis accounted for about 62 million people living in the 246 catalog units characterized by improvements in dissolved oxygen. The population represented in the "before and after" analysis thus accounts for about 31 percent of the total continental US population of 197 million recorded in the 1990 census. During both the "before" and "after" periods, "dry" hydrologic conditions (< 75 percent summer mean streamflow) were recorded in about 90 percent of all catalog units for at least 1 of the 5 years of record with "dry" conditions persisting for an average of 2.5 years during 1961-1965 and 2.7 years during 1986-1990. Filters used in the "before and after" analysis included specification of hydrologic conditions as "dry,"" normal," or "wet" and selection of RF1 reaches defined by the impact of nonpoint source and point source pollutant loads as "only PS impacted," "only NPS impacted," or "PS and NPS impacted." The results presented in the report emphasized the worst-case findings for "only PS impacted" reaches under "dry" hydrologic conditions. The complete listing of the "before and after" data for dissolved oxygen and BOD_5 for the 246 catalog units with "before and after" data is presented in Appendix D of this report. "Before and after" BOD_5 data were correlated with "before and after" dissolved oxygen data in the discussion of catalog unit and RF1 reach scale trends for the Upper White River catalog unit in Indiana. The data tables identify the waterways ranked by the "before and after" change in oxygen and BOD_5 from the greatest improvement to the worst degradation.

In response to a reviewer's suggestion, the "before and after" analysis was performed for RF1 reaches impacted only by NPS loads under "dry" flow conditions. The results indicated that (a) very little "before and after" data was available to characterize trends for reaches impacted only by NPS and (b) oxygen conditions

in NPS reaches were not characterized by the significant improvements identified for PS impacted reaches. Selection of "dry" hydrologic conditions for at least 1 of 5 years for "only NPS impacted" reaches resulted in extraction of a small sample of only 37 catalog units with sufficient "before and after" data sets. The *greatest* improvement was characterized by oxygen increasing from 3.0 mg/L during 1961-1965 to 6.2 mg/L during 1986-1990 in the Lower Dan River watershed in North Carolina.

Issue 7: Statistical Methods

Reviewer:

Smith

Comment summary:

DO deficit data should be used instead of DO concentration data to account for possible differences in temperature during the before and after periods. Discussions of the statistical effect of data aggregation should be clarified, It should be stressed that pooling reach data at the national level enhances the ability to distinguish signal from noise and obtain statistically significant results. More discussion is needed on uncertainty, particularly where results are based on limited data. Some assessment should be made of the degree of statistical dependence of DO concentration between reach-level results, to ensure there is no bias due to "clustering" of monitoring stations in adjacent or nearby reaches.

Response:

Discussion has been added in Chapter 3 to address the selection of DO concentration over DO saturation or DO deficit (see also Issue #4). The authors have adopted the wording "data aggregation" in conjunction with "spatial scale" in the description of the statistical methods and results. Discussion has also been included in Chapter 3 to address uncertainty. Since the study authors found similar results in before versus after DO concentration changes at all three data aggregation scales (i.e., two-thirds of hydrologic units at the reach, catalog unit, and major river basin scales), additional analysis on spatial clustering of monitoring stations was not performed.

Issue 8: Geographic Representation of Case Studies

Reviewers:

Lagnese, Parker, Kirk, Okun

Comment summary:

A few reviewers pointed out the lack of geographic diversity of case study sites. At least one case study should have been selected to represent the western arid region. Other reviewers commented that the case studies really presented a good documentation of improvements in water quality although how these were related to the CWA was not always clearly explained.

Response:

The discussion on the study authors' logic for case study site selection has been expanded in Chapter 4. Geographic diversity was not a priority in the selection process; instead, selection focused on heavily urbanized, highly populated waterways with a history of water quality problems related to municipal wastewater treatment discharges. The case study chapters have been expanded to clarify the relationship of observed water quality conditions to actions related to the CWA.

Issue 9: Cost versus Benefit Analysis

Reviewers:

Lagnese, Parker, Linsky, Okun

Comment summary:

Not enough explanation is provided of how cost and benefit data were compiled for the three case studies (Potomac, Upper Mississippi, and Willamette Rivers) presented in the Executive Summary and Chapter 4 (overview of cases). A cost/benefit analysis is a major issue for the evaluation of the effectiveness of the CWA. The cost/benefit data were, in fact, knowingly included in the Executive Summary without any explanation at all.

Response:

The study authors agree that cost/benefit is a major issue for evaluation of the effectiveness of the CWA. The original scope of this study was never intended to include cost versus benefit analysis of pollution controls, but rather was to focus on national trends in population, wastewater treatment plant design capacity, and water quality conditions in receiving waters. The cost versus benefit data provided for the three case study sites was prepared under a separate study and does not extend to the other six case studies or nationally. The study authors have chosen to remove the cost versus benefits data and information completely from the document to avoid confusion or presentation of incomplete analyses since this is a national-scale study.

Issue 10: Editing for final report

Reviewers:

Wolman, Landman, Kirk

Comment summary:

Eliminate all self-congratulatory phrases claiming “success” or what a wonderful methodology and study this is, etc. Edit out redundant material presented in several chapters. The tone of the report is “biased” and “not objective.” Provide a clear summery and conclusions sections where appropriate throughout the document.

Response:

The document has been extensively revised and rewritten since the peer review. The study authors have attempted to remove wherever possible any redundancy or biased statements. A summary and conclusions section has been added at the end of each chapter. The entire study is based on factual information that speaks for itself, allowing the readers to draw their own conclusions.